Alexandru Baltag
Nina Gierasimczuk🄳
Sonja Smets

Truth-Tracking by Belief Revision

Abstract. We study the learning power of iterated belief revision methods. Successful learning is understood as convergence to correct, i.e., true, beliefs. We focus on the issue of universality: whether or not a particular belief revision method is able to learn everything that in principle is learnable. We provide a general framework for interpreting belief revision policies as learning methods. We focus on three popular cases: conditioning, lexicographic revision, and minimal revision. Our main result is that conditioning and lexicographic revision can drive a universal learning mechanism, provided that the observations include only and all true data, and provided that a non-standard, i.e., non-well-founded prior plausibility relation is allowed. We show that a standard, i.e., well-founded belief revision setting is in general too narrow to guarantee universality of any learning method based on belief revision. We also show that minimal revision is not universal. Finally, we consider situations in which observational errors (false observations) may occur. Given a fairness condition, which says that only finitely many errors occur, and that every error is eventually corrected, we show that lexicographic revision is still universal in this setting, while the other two methods are not.

Keywords: Belief revision, Dynamic Epistemic Logic, Formal learning theory, Truth-tracking.

Introduction

At the basis of the modeling of intelligent behavior lies the idea that agents integrate new information into their prior beliefs and knowledge. Intelligent agents are assumed to be endowed with some learning methods, which allow them to change their beliefs on the basis of assessing new information. But how effective is an agent's learning method in eventually finding the truth? To make this question precise and to answer it, we borrow concepts from formal learning theory and adapt them to the commonly used model of beliefs, knowledge, and belief change, namely that of possible worlds.

A set S of possible worlds, which we will call the *state space*, together with a family $\mathcal{O}$ of observable propositions, represents the agent's epistemic

space, or knowledge state. Note that the sets S and $\mathcal{O}$ do not have to be finite, but are assumed to be at most countable. Intuitively, the epistemic space represents the uncertainty range of the agent. She can consider some possible worlds to be more plausible than others. This is represented by a total preorder on possible worlds, called a *plausibility preorder*. It captures the agent's assessments concerning which of any two worlds s, s' is more likely to be the actual one. Such an assessment can obviously be based on many different factors, in particular on the assessed level of simplicity, or on consistency with previous observations. An epistemic space together with a plausibility preorder is called a *plausibility space*.

In order to represent the dynamic aspects of knowledge and belief we will use belief revision methods, which, triggered by incoming information, change the epistemic (plausibility) space. The change can occur through removal of the states incompatible with the new information, or through a revision of the plausibility relation. Many belief revision policies proposed in the literature are formulated, or can be reconstructed, within our setting. In this paper we investigate three basic policies: conditioning, lexicographic revision [38,39], and minimal revision [18]. The goal is to see how they can be viewed as learning methods, and to investigate their learning power, i.e., the ability to identify the real world from the incoming information.

We obtain our results by defining learning methods based on belief revision policies. We show that learning from positive data via conditioning and lexicographic revision is universal, i.e., that those learning methods can uniformly identify the real world in the limit, when starting in any epistemic space in which the real world is identifiable (via any learning method). However, that happens only if the agent's prior plans/dispositions for belief revision are suitably chosen; and not every such prior is suitable. Furthermore, we show that the most conservative belief revision method, minimal revision, is not universal.

Our approach brings together methods of formal learning theory [FLT, see, e.g., 37] and Dynamic Epistemic Logic [DEL, see 7,8,14,22]. The interest in bringing together learning theory and belief revision theory has appeared before within at least two lines of research. Firstly, in [30–34] some classical belief revision policies were treated as learning strategies. Secondly, in [35,36] the connection has been rooted in the classical AGM framework [1]. Finally, in [23–27] the set learning paradigm (also called *language* (or *set*) *learning*) has been connected with epistemic and doxastic logics of belief revision [3,9–11,13,21]. The present paper is a continuation of the latter line of research, and is in fact a thorough presentation of results announced in a previously published extended abstract [5].

We are chiefly concerned with the counterpart of one of the central notions in formal learning theory, namely *identifiability in the limit* [28]. Hence, we focus on *stabilizing* to a correct *belief.*[1] We hence investigate the *reliability* of mind-changing strategies, i.e., the possibility of converging to an accurate hypothesis after a finite number of mind-changes.

1. Notation and Basic Definitions

The agent is represented by her epistemic space, i.e., a range of possible worlds that satisfy relevant observables.

DEFINITION 1. Let S be set of possible worlds and let $\mathcal{O} \subseteq \mathcal{P}(S)$ be a set of observable propositions (also called 'observables'). We will assume that both sets are possibly infinite, but at most countable, and that there are no two possible worlds that make exactly the same propositions true. The pair $\mathbb{S} = (S, \mathcal{O})$ is then called an *epistemic space*.

The epistemic space represents an agent who does not favor any possibility over any other. We break that symmetry by introducing a *plausibility relation*, $\preceq$.[2]

DEFINITION 2. Let $\mathbb{S} = (S, \mathcal{O})$ be an epistemic space, and $\preceq \subseteq S \times S$ be a total preorder.[3] The structure $\mathbb{B}_\mathbb{S} = (S, \mathcal{O}, \preceq)$ is called a *plausibility space*.

Since we allow for the epistemic space to be infinite, the question of well-foundedness of the plausibility space becomes very relevant. We do not restrict our considerations to well-founded spaces, which we call *standard plausibility spaces* due to their popularity in their belief revision literature.

DEFINITION 3. A *standard plausibility space* $\mathbb{B}_\mathbb{S} = (S, \mathcal{O}, \preceq)$ is one whose plausibility relation $\preceq$ is well-founded (i.e., there is no infinite descending chain $s_0 \succ s_1 \succ \ldots \succ s_n \succ \ldots$, where $\prec$ is the strict plausibility relation, given by: $s \preceq t$ and $t \not\preceq s$).

---

[1]The emergence of the stronger epistemic state of irrevocable knowledge can be linked to a more restrictive kind of identifiability, finite identifiability [see 19, 20, 27]. It has also been recently investigated in the context of action learning in dynamic epistemic logic [16, 17].

[2]Note that we interpret $s \preceq t$ as 's is at least as plausible as t'.

[3]In other words, the binary relation $\preceq$ is total, reflexive, and transitive in S.

1.1.   Knowledge and Belief in Epistemic Spaces

Let us briefly discuss the interpretation of knowledge and belief in our setting. Statements about knowledge and belief are not taken to be observables, but are meaningful within the epistemic spaces. Let us take a plausibility space $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$. Given a proposition $p \subseteq S$, we say that the agent knows that p, Kp, if and only if $S \subseteq p$. In other words, knowledge is a global modality—we say that the agent knows p iff p is true in all possible worlds of the plausibility space $\mathbb{B}_{\mathbb{S}}$. In such case we will write $\mathbb{B}_{\mathbb{S}} \models Kp$.

As we mentioned before, most of the epistemic doxastic logic and belief revision literature deals with standard, i.e., well-founded plausibility structures. The well-foundedness assumption has at least two advantages. Firstly, it allows us to canonically assign ordinal numbers to states [so-called *Spohn ordinals* or 'degrees of implausibility', see 39]. Secondly, it leads to a simple definition of belief, which can then be understood as 'truth in all the most plausible worlds'. In any standard plausibility space $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, the agent believes p, Bp, if and only if $\min_{\preceq} S \subseteq p$, where for any set $X \subseteq S$, $\min_{\preceq} X$ is the set of all most plausible worlds in X, defined as $\{t \in X \mid t \preceq s \text{ for all } s \in X\}$.[4] For simplicity, given any $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, we will write $\min \mathbb{B}_{\mathbb{S}}$, meaning $\min_{\preceq} S$. Obviously, this definition is equivalent to $\min \mathbb{B}_{\mathbb{S}} \subseteq p$ in well-founded plausibility spaces.

We do not assume well-foundedness because, as we will show, it prevents one from learning what one could have learned with a non-well-founded plausibility relation. To make the belief operator meaningful also in the non-well-founded cases, we will say that the agent believes that p, Bp, if and only if $\exists w \, \forall u \preceq w \, u \in p$. In such case we will write $\mathbb{B}_{\mathbb{S}} \models Bp$.

1.2.   Observable Propositions

An observable proposition is identified with the set of those possible worlds which makes it true. These propositions can be empirically encountered (observed) by an agent and hence can be viewed as *data* or *evidence* for learning: an agent can witness them. This does not mean that they are all observable at the same time. We will assume that at each step of learning consists of only one observation.[5]

---

[4]It is easy to see that, if $\preceq$ is well-founded, then $\min_{\preceq} X \neq \emptyset$ whenever $X \neq \emptyset$.

[5]In general, the set of observables can be closed under certain operations, e.g., under negation (if 'negative data' are observed), or under finite intersection (if 'conjunctions' are observed). Under the usual possible-world interpretation, $\mathcal{O}$ can be viewed as the set of (atomic) propositions, or, if the stress is put on closure under certain operations (e.g., negation or conjunction), as a set encoding the valuation for a relevant logical language.

The agent revises her beliefs (plausibility space) in response to the observations. We assume that the agent is inductively presented with an endless stream of observables.

DEFINITION 4. Let $\mathcal{S} = (S, \mathcal{O})$ be an epistemic space.
A *data stream* is an infinite sequence $\vec{O} = (O_0, O_1, \ldots)$ of data $O_i \in \mathcal{O}, i \in \mathbb{N}$.
A *data sequence* is a finite sequence $\sigma = (O_0, \ldots, O_n)$.

The intuition behind the streams of data is that at stage i, the agent observes the information in O_i. A data stream captures a possible future history of observations in its entirety, while a data sequence captures only a finite part of such a history.

DEFINITION 5. Let $\vec{O} = (O_0, O_1, \ldots)$ be a data stream, and let $\sigma = (\sigma_0, \ldots, \sigma_n)$ be a data sequence.
$\vec{O}_n$ stands for the n-th observation in $\vec{O}$.
$\vec{O}[n]$ stands for the the initial segment of $\vec{O}$ of length n, $(O_0, \ldots, O_{n-1})$.
$\mathrm{set}(\vec{O}) := \{O \mid O \text{ is an element of } \vec{O}\}$ stands for the set of all data occuring in $\vec{O}$; we similarly define $\mathrm{set}(\sigma)$, where σ is a finite data sequence.
$\sigma * \vec{O} := (\sigma_0, \ldots, \sigma_n, O_0, O_2, \ldots)$ is the concatenation of the finite sequence σ with the infinite stream $\vec{O}$.

A data stream is *sound* iff it presents only true observables.

DEFINITION 6. A data stream $\vec{O}$ is *sound with respect to state s* iff every element listed in $\vec{O}$ is true in s, i.e., $s \in \vec{O}_n$ for all $n \in \mathbb{N}$.

A data stream is *complete* iff it presents every true observable property.

DEFINITION 7. A data stream $\vec{O}$ is *complete with respect to state s* iff every observable true in s is listed in $\vec{O}$, i.e., for any $O \in \mathcal{O}$, if $s \in O$ then $O = \vec{O}_n$ for some $n \in \mathbb{N}$.

Let $\mathcal{O}_s$ stand for the set of all observables that are true in s, i.e., $\mathcal{O}_s = \{O \in \mathcal{O} \mid s \in O\}$. A data stream $\vec{O}$ is then *sound and complete with respect to state s* if and only if $\mathcal{O}_s = \mathrm{set}(\vec{O})$.

In most of this paper we assume the data streams to be sound and complete with respect to the actual world, i.e., all observed data is true, and all true data will eventually be observed.[6] Those are the most favorable conditions for learning (in the limit) the whole truth about the identity of the actual world. Nonetheless, learning from such data is not trivial, as will be seen below.

---

[6] In computational learning theory, such sound and complete data streams are called 'texts' [29] or 'environments' [36].

2.   Learning and Belief Revision Methods

In this section, we introduce a formal framework for belief change, learning functions, and belief revision operators. Depending on the epistemic space, a *learning function* responds to a given data sequence with a new conjecture.

DEFINITION 8. Let $\mathbb{S} = (S, \mathcal{O})$ be an epistemic space and let σ be a data sequence. A *learning method* is a function L that on the input of $\mathbb{S}$ and data sequence σ outputs some set of worlds $L(\mathbb{S}, \sigma) \subseteq S$, called a *conjecture.*

Learning methods can have various properties; for instance, the learner can be forgetful or conservative in revising her conjectures. Below we list several properties of this type.

DEFINITION 9. A learning method L is:

(1) *weakly data-retentive* iff $L(\mathbb{S}, \sigma) \neq \emptyset$ implies $L(\mathbb{S}, \sigma) \subseteq \sigma_n$, for any $\mathbb{S}$ and $\sigma = (\sigma_0, \ldots, \sigma_n)$;

(2) *strongly data-retentive* iff $L(\mathbb{S}, \sigma) \neq \emptyset$ implies $L(\mathbb{S}, \sigma) \subseteq \bigcap_{i \in \{0, \ldots, n\}} \sigma_i$, for any $\mathbb{S}$ and $\sigma = (\sigma_0, \ldots, \sigma_n)$;

(3) *conservative* iff $\emptyset \neq L(\mathbb{S}, \sigma) \subseteq p$ implies $L(\mathbb{S}, \sigma) = L(\mathbb{S}, \sigma * p)$, for any $\mathbb{S}$, σ, and $p \in \mathcal{O}$;

(4) *data-driven* if it is both conservative and weakly data-retentive;

(5) *memory-free* iff $L(\mathbb{S}, \sigma) = L(\mathbb{S}, \sigma')$ implies $L(\mathbb{S}, \sigma * p) = L(\mathbb{S}, \sigma' * p)$, for any $\mathbb{S}$ and every two data sequences σ, σ', and every $p \in \mathcal{O}$.

For the reader familiar with belief revision theory, let us now briefly compare the above properties with the well-known AGM postulates [1]. Weak data retention means that the current conjecture always entails the most recently observed data. If we interpret conjectures as beliefs, this intuitively corresponds to the AGM *Success Postulate.* Strong data retention says that the current conjecture always accounts for all data that have been encountered till now. Between those two extremes, a learner with bounded memory could retain some fixed finite amount of data. Conservativity requires that the agent sustains the same conjecture whenever the new piece of data is already entailed by her old conjecture. A learning method is memory-free if, at each stage, the new conjecture depends only on the previous conjecture and the new datum.[7] As we will see later, this assumption poses severe

---

[7] The latter property was the original intention behind the AGM notation $T * \varphi$, for revision of a theory T with a new piece of data φ [see 1].

problems for *iterated* belief revision. In fact, some standard belief revision methods implementing the AGM postulates are not memory-free: the new belief depends in addition on some hidden parameter, namely the old plausibility relation.

We now turn to belief revision methods. In our logical setting, they are transformations of plausibility spaces triggered by the incoming data.

DEFINITION 10. A *one-step revision method* is a function R_1 that, for any plausibility space $\mathbb{B}_\mathbb{S}$ and any observable $p \in \mathcal{O}$, outputs a new plausibility space $R_1(\mathbb{B}_\mathbb{S}, p)$.

A *(iterated) belief revision method* R is obtained by iterating a one-step revision method R_1:

$$R(\mathbb{B}_\mathbb{S}, \lambda) = \mathbb{B}_\mathbb{S}, \text{ for the empty data sequence } \lambda;$$
$$R(\mathbb{B}_\mathbb{S}, \sigma * p) = R_1(R(\mathbb{B}_\mathbb{S}, \sigma), p).$$

Now we show how to define learning functions in terms of belief revision. First, one must assign an *initial plausibility order* to each epistemic space to turn it into a plausibility space.

DEFINITION 11. Let $\mathbb{S} = (S, \mathcal{O})$ be any epistemic space. A *prior plausibility assignment* PLAUS is a map that assigns to $\mathbb{S}$ some plausibility relation $\preceq$ on S, thus converting it into a plausibility space PLAUS$(\mathbb{S}) = (S, \mathcal{O}, \preceq)$.

Then one can apply the belief revision operation to each new datum and recover the resulting belief state as the set of all worlds that are minimal (most plausible) in the revised plausibility space.

DEFINITION 12. Every belief revision method R, together with a prior plausibility assignment PLAUS, generates a learning method L_R^{PLAUS}, called a *belief revision-based learning method*, given by:

$$L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) := \min R(\text{PLAUS}(\mathbb{S}), \sigma).$$

The previously defined properties of learning functions (Definition 9) can be now applied to belief revision methods:

DEFINITION 13. A belief revision method R is called *weakly data-retentive (strongly data-retentive, conservative*, or *data-driven)* iff for any prior plausibility assignment PLAUS, the induced learning method L_R^{PLAUS} is weakly data-retentive (strongly data-retentive, conservative, or data-driven).

The properties that belief revision methods inherit from their corresponding learning methods can be characterised in terms of belief.

PROPOSITION 1. *Let R be a belief revision method.*

(1) R *is weakly data-retentive iff* $R(\mathbb{B}_{\mathbb{S}}, \sigma) \models B\sigma_n$, *for any* $\mathbb{B}_{\mathbb{S}}$ *and* $\sigma = (\sigma_0, \ldots, \sigma_n)$;

(2) R *is strongly data-retentive iff* $R(\mathbb{B}_{\mathbb{S}}, \sigma) \models B\sigma_i$, *for any* $\mathbb{B}_{\mathbb{S}}$, $\sigma = (\sigma_0, \ldots, \sigma_n)$, *and any* $i \in \{0, \ldots, n\}$;

(3) *If R is conservative then* $R(\mathbb{B}_{\mathbb{S}}, \sigma) \models Bq$ *iff* $R(\mathbb{B}_{\mathbb{S}}, \sigma * p) \models Bq$, *for any* $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, σ, $q \subseteq S$, *and* $p \in \mathcal{O}$, *such that* $R(\mathbb{B}_{\mathbb{S}}, \sigma) \models Bp$.

PROOF. The left-to-right implications of (1) and (2) are trivial, given the semantics of the belief operator. For the right-to-left implication of (1), let R be a belief revision method, let $\mathbb{S} = (S, \mathcal{O})$, PLAUS be any plausibility assignment, and $\sigma = (\sigma_0, \ldots, \sigma_n)$. Assume that $R(\text{PLAUS}(\mathbb{S}), \sigma) \models B\sigma_n$, and that that $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) \neq \emptyset$. Then $\min R(\text{PLAUS}(\mathbb{S}), \sigma) \subseteq \sigma_n$, by hypothesis. Since $\min R(\text{PLAUS}(\mathbb{S}), \sigma) = L_R^{\text{PLAUS}}(\mathbb{S}, \sigma)$, $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) \subseteq \sigma_n$. The proof of the right-to-left implication in the second assertion is analogous.

For (3), let R be a conservative belief revision method, i.e., L_R^{PLAUS} is conservative, for any PLAUS. Let us take a PLAUS, $\mathbb{S}$, σ, and $p \in \mathcal{O}$, such that $\emptyset \neq L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) \subseteq p$, and so $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) = L_R^{\text{PLAUS}}(\mathbb{S}, \sigma * p)$. It suffices to show that $R(\text{PLAUS}(\mathbb{S}), \sigma) \models Bq$ iff $R(\text{PLAUS}(\mathbb{S}), \sigma * p) \models Bq$, for any $q \subseteq S$.

For the left to right direction, assume that $R(\text{PLAUS}(\mathbb{S}), \sigma) \models Bq$. Since $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) \neq \emptyset$, it follows by the definition of B that $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) \subseteq q$. Then, since $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma) = L_R^{\text{PLAUS}}(\mathbb{S}, \sigma * p)$, the definition of belief revision based learning method yields that $L_R^{\text{PLAUS}}(\mathbb{S}, \sigma * p) = \min R(\text{PLAUS}(\mathbb{S}), \sigma * p) \subseteq q$, which by the semantics of B implies that $R(\text{PLAUS}(\mathbb{S}), \sigma * p) \models Bq$. For the converse the argument is analogous. ∎

We also define additional, specific to belief revision properties: strong conservativity and history independence.

DEFINITION 14. A belief revision method R is:

(a) *strongly conservative* iff for any $\mathbb{B}_{\mathbb{S}}$, σ, and $p \in \mathcal{O}$ such that $R(\mathbb{B}_{\mathbb{S}}, \sigma) \models Bp$, $R(\mathbb{B}_{\mathbb{S}}, \sigma) = R(\mathbb{B}_{\mathbb{S}}, \sigma * p)$, i.e., if the new piece of data was already believed, R does not change the plausibility space at all.

(b) *history-independent* iff for any $\mathbb{B}_{\mathbb{S}}$, $p \in \mathcal{O}$, and any data sequences σ, π we have that if $R(\mathbb{B}_{\mathbb{S}}, \sigma) = R(\mathbb{B}_{\mathbb{S}}, \pi)$ then $R(\mathbb{B}_{\mathbb{S}}, \sigma * p) = R(\mathbb{B}_{\mathbb{S}}, \pi * p)$, i.e., R's output at any stage depends only on the previous output and the most recently observed data.

Strongly conservative belief revision methods not only keep the old conjecture the same (as in the case of conservative learning methods), but, when

receiving truthful information, they do not change anything within the plausibility space. As we will see below, the classical belief revision methods are not necessarily strongly conservative. However, every iterated one-step belief revision method must be history-independent. History-independent methods do not require that the agent retains all past events: only the last plausibility space and the new datum are enough to determine the next plausibility space. However, as we will show in the next section, the corresponding learning is not necessarily memory-free.

3.   Some Iterated Belief Revision Methods

Below we consider three basic belief revision methods that received considerable attention in dynamic epistemic logic and belief revision theory.

Conditioning. First we focus on the revision by *conditioning* [38,39], also called *update* in DEL [11,13]. This method operates by deleting those worlds that do not satisfy the newly observed data.

DEFINITION 15. *Conditioning*, Cond, is an (iterated) belief revision method based on the one-step belief revision method Cond_1 that takes $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$ and $p \in \mathcal{O}$, and outputs a restriction of $\mathbb{B}_{\mathbb{S}}$ to p, i.e., $\text{Cond}_1(\mathbb{B}_{\mathbb{S}}, p) = (S^p, \mathcal{O}, \preceq^p)$, where $S^p = S \cap p$, and $\preceq^p = \preceq \cap (S^p \times S^p)$.

Since conditioning throws out all worlds inconsistent with current information, it is easy to see that:

PROPOSITION 2. Cond *is strongly data-retentive on sound data streams.*

PROOF. Let us take any $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$ and a $\sigma = (\sigma_0, \ldots, \sigma_n)$ sound with respect to an $s \in S$. By Proposition 1, it suffices to show that $\text{Cond}(\mathbb{B}_{\mathbb{S}}, \sigma) \models B\sigma_i$ for $i \in \{0, \ldots, n\}$. Observe that indeed, for all w such that $w \preceq s$, $\text{Cond}(\mathbb{B}_{\mathbb{S}}, \sigma) \models B\sigma_i$, for $i \in \{0, \ldots, n\}$. It is the case because σ is sound with respect to σ, and since all worlds inconsistent with elements of σ have already been eliminated by Cond. ∎

While conditioning is obviously conservative, it is not strongly conservative, since new information can rule out some worlds without ruling out all of the most plausible worlds:

PROPOSITION 3. Cond *is not strongly conservative.*

PROOF. Consider $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, with $\mathcal{O} = \{p, q\}$, $S = \{s, t\}$ such that $p = \{s, t\}$, $q = \{s\}$, and $s \prec t$ (see Figure 1). Then $\mathbb{B}_{\mathbb{S}} \models Bq$, but $\text{Cond}(\mathbb{B}_{\mathbb{S}}, \lambda) \neq \text{Cond}(\mathbb{B}_{\mathbb{S}}, \lambda * q)$. ∎
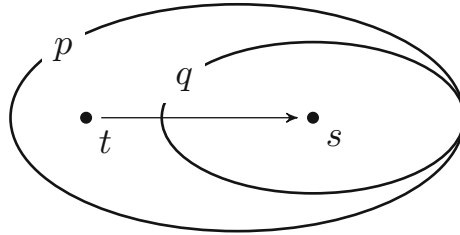
Figure 1. Plausibility space from the proof of Proposition 3. The arrow points to the more plausible world, in this case $s \prec t$

Lexicographic Revision. Lexicographic revision [38,39], also known as *radical upgrade* in Dynamic Epistemic Logic [11,13] does not delete any worlds. Instead, it 'promotes' all the worlds satisfying the new piece of data, making them more plausible than all the worlds that do not satisfy it; while within the two zones, the old order is kept the same.

DEFINITION 16. *Lexicographic revision*, Lex, is a belief revision method based on the one-step belief revision method Lex_1 that takes $\mathbb{B}_{\mathbb{S}}$ and $p \in \mathcal{O}$, and outputs a revised plausibility space: $\text{Lex}_1(\mathbb{B}_{\mathbb{S}}, p) = (S, \mathcal{O}, \preceq')$, where for all $t, w \in S$, $t \preceq' w$ iff ($t \preceq_p w$ or $t \preceq_{\bar{p}} w$ or ($t \in p \wedge w \notin p$)), where: $\preceq_p = \preceq \cap (p \times p)$, $\preceq_{\bar{p}} = \preceq \cap (\bar{p} \times \bar{p})$, and $\bar{p}$ stands for the complement of p in S.

It is easy to see that Lex is weakly data-retentive and conservative. However, it does not satisfy the strong versions of these properties:

PROPOSITION 4. Lex *is not strongly data-retentive on arbitrary streams.*

PROOF. The proof is analogous to that of Proposition 2.   ■

PROPOSITION 5. Lex *is strongly data-retentive on sound streams.*

PROOF. Let us take $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$. Let us also fix $s \in S$ and assume that $\sigma = (\sigma_0, \ldots, \sigma_n)$, is sound with respect to s, i.e., $\mathbb{B}_{\mathbb{S}}, s \models \bigwedge \text{set}(\sigma)$. It is easy to see that $\min \text{Lex}(\mathbb{B}_{\mathbb{S}}, \sigma) = \sigma_0 \cap \ldots \cap \sigma_n \neq \emptyset$. Hence, $\mathbb{B}_{\mathbb{S}} \models B(\bigwedge \text{set}(\sigma))$.
                                                                                          ■

PROPOSITION 6. *Lexicographic revision is not strongly conservative.*

PROOF. Let us consider a plausibility space in Figure 2: $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$, where $S = \{s, t, u\}$, $\mathcal{O} = \{p, q\}$, and $p = \{s, u\}$, $q = \{s, t\}$. Moreover, the plausibility $\preceq$ gives the following order: $s \prec t \prec u$, and that $\sigma = (p)$. Clearly, $\mathbb{B}_{\mathbb{S}} \models Bp$. However, after receiving $\sigma_0 = p$, the revision method will still put world u to be more plausible than t, and therefore $\mathbb{B}_{\mathbb{S}} \neq \text{Lex}(\mathbb{B}_{\mathbb{S}}, p)$.   ■
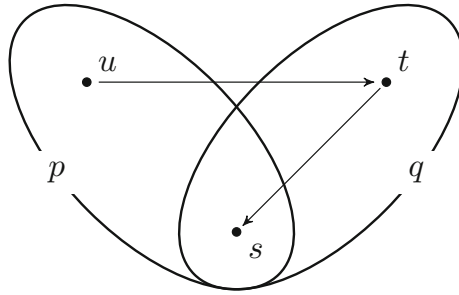
Figure 2. Plausibility space from the proof of Proposition 6

PROPOSITION 7. *A learning method generated from a history-independent belief revision method does not have to be memory-free.*

PROOF. We prove this proposition by showing an example, a belief revision method R that is history-independent but the learning method that it induces is not memory-free (see Figure 3). Let R be Lex. Lex is clearly history-independent, because it is an iterated one-step revision method. To see that $L_{\texttt{Lex}}$ is not memory-free consider the following plausibility space $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$ with $\mathcal{O} = \{p, q, r\}$ and $S = \{s, t, u\}$, such that $p = \{s, u\}$, $q = \{t, u\}$, $r = \{s, t, u\}$. Take $\sigma = (p)$ and $\sigma' = (r)$:

(1) $\texttt{Lex}(\mathbb{B}_{\mathbb{S}}, \sigma)$ gives the plausibility order: $s \prec u \prec t$;

(2) $\texttt{Lex}(\mathbb{B}_{\mathbb{S}}, \sigma')$ gives the plausibility order: $s \prec t \prec u$.

So, $L_{\texttt{Lex}}(\mathbb{S}, \sigma) = L_{\texttt{Lex}}(\mathbb{S}, \sigma') = \{s\}$. Assume now that the next observation is q. Then clearly $L_{\texttt{Lex}}(\mathbb{S}, \sigma * q) = \{u\}$, while $L_{\texttt{Lex}}(\mathbb{S}, \sigma' * q) = \{t\}$. Therefore, for the belief revision method Lex there is a $p \in \mathcal{O}$ such that: $L_{\texttt{Lex}}(\mathbb{S}, \sigma) = L_{\texttt{Lex}}(\mathbb{S}, \sigma')$ but $L_{\texttt{Lex}}(\mathbb{S}, \sigma * p) \neq L_{\texttt{Lex}}(\mathbb{S}, \sigma' * p)$. So, $L_{\texttt{Lex}}$ is not memory-free.
∎

Minimal Revision. The minimal revision method [18,38], known as *conservative upgrade* in DEL [11,13], is 'conservative' in the sense that it keeps as much as possible of the old structure. More precisely, the most plausible states satisfying the new piece of data become the most plausible overall; while in the rest of the space, the old order is kept the same.

DEFINITION 17. *Minimal revision* Mini, is a belief revision method based on the one-step belief revision method $\texttt{Mini}_1$ that takes a plausibility space $\mathbb{B}_{\mathbb{S}}$ and a proposition $p \in \mathcal{O}$, and outputs a new plausibility space in the
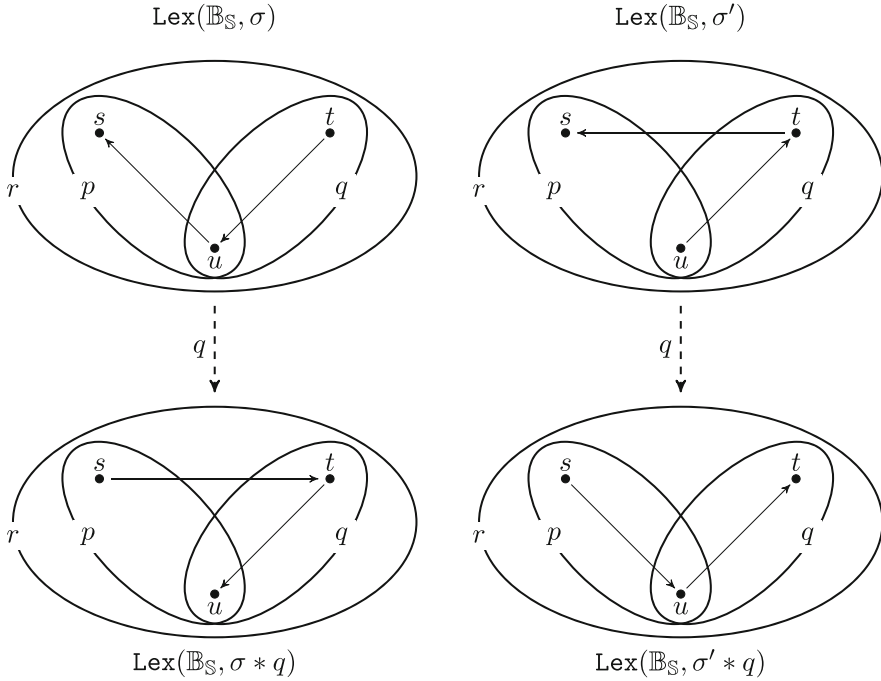
Figure 3. The transformations of the plausibility space from the proof of Proposition 7

following way: $\mathtt{Mini}(\mathbb{B}_\mathbb{S}, p) = \mathbb{B}_\mathbb{S}' = (S, \mathcal{O}, \preceq')$ where for all $t, w \in S$, if $t \in min_{\preceq}p$ and $w \notin min_{\preceq}p$, then $t \preceq' w$, otherwise $t \preceq' w$ iff $t \preceq w$.

Minimal revision is obviously weakly data-retentive—it leads to a belief that accounts for the last datum. However, it does not retain more than that. Consider the plausibility space in Figure 4, and the sequence $\sigma = (p, q)$ (which is sound with respect to world u). After receiving p the plausibility order becomes $s \prec t \prec u$. Then q comes in and now our method gives the order $t \prec s \prec u$. So p is no longer believed after q was given. Therefore:

PROPOSITION 8. *Minimal revision is not strongly data-retentive.*

Moreover, the minimal revision is conservative: as long as the incoming information is already believed, beliefs do not change, since the minimal worlds do not change. In this case we can say even more: nothing about the plausibility space changes, and so:

PROPOSITION 9. *Minimal revision is strongly conservative.*

The properties introduced in this section capture some interesting distinctions between belief revision methods. While conditioning and lexicographic
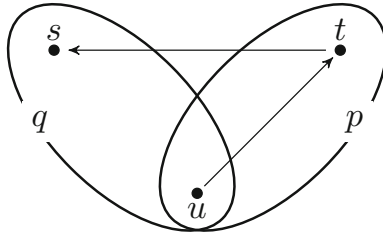
Figure 4. Example of a plausibility space

revision are quite similar, differing only with respect to their strong retention capacity, minimal revision is different in two respects. It is not strongly data retentive, even on sound data streams. In the next section we will see that this combination of properties negatively affects learning [similar observations about the interaction of those properties were given in 30,31].

4. Convergence to Truth

Formal learning theory is concerned with *reliable* learning methods, i.e., those that can be relied upon (when observing a sound and complete data stream) to find the complete truth in the limit no matter what the actual world is, as long as it is among the possibilities allowed by the initial epistemic space $\mathbb{S}$.[8] Following the learning-theoretic terminology, we say in this case that the real world has been *identified in the limit*.

DEFINITION 18. Given an epistemic space $\mathbb{S} = (S, \mathcal{O})$, a world $s \in S$ is *identified in the limit by a learning method* L if, for every sound and complete data stream for s, there exists a finite stage after which L outputs the singleton $\{s\}$ from then on.

We say that the epistemic space $\mathbb{S}$ is *identified in the limit by* L iff all its worlds are identified in the limit by L.

An epistemic space $\mathbb{S}$ is *identifiable in the limit* (*learnable*) if there exists a learning method L that can identify it in the limit.

Learning methods differ in their learning power. We are interested in the *most powerful* among them, those that are *universal*—they can learn any epistemic space that is learnable.

---

[8]For a discussion of reliability of belief revision methods see [34].

DEFINITION 19. A learning method L is *universal on a class $\mathcal{C}$ of epistemic spaces* if it can identify in the limit every epistemic space in $\mathcal{C}$ that is identifiable in the limit. A *universal learning method* is one that is universal on the class of all epistemic spaces.

In the remainder of this paper we focus on learning methods that are generated by iterated belief revision methods. For brevity, we will attribute the ability of identification in the limit also to belief revision policies.

DEFINITION 20. An epistemic space $\mathbb{S}$ is *identified in the limit by a belief revision method R* if there exists a prior plausibility assignment PLAUS such that the generated learning method L_R^{PLAUS} identifies S in the limit.
The epistemic space $\mathbb{S}$ is *standardly identified in the limit* by R if there exists a well-founded prior plausibility assignment PLAUS such that L_R^{PLAUS} identifies S in the limit.

DEFINITION 21. A revision method R is *universal on a class $\mathcal{C}$ of epistemic spaces* if it can identify in the limit every epistemic space in $\mathcal{C}$ that is identifiable in the limit.
R is *standardly universal on a class $\mathcal{C}$* if it can standardly identify in the limit every epistemic space in $\mathcal{C}$ that is identifiable in the limit.

Our main result is the existence of AGM-like universal belief revision methods. The main technical difficulty of this part is the construction of an appropriate prior plausibility order. To define it, we use the concept of locking sequences introduced in [15] and that of finite tell-tale sets proposed in [2]. We adjust the classical notion of finite tell-tales, and use it in the construction of a suitable prior plausibility assignment that, together with conditioning and lexicographic revision, generates universal learning methods.

The first observation is that if convergence occurs, then there is a finite sequence of data that 'locks' the corresponding sequence of conjectures on a correct answer. This finite sequence is called a *locking sequence* [15].

DEFINITION 22. Let an epistemic space $\mathbb{S} = (S, \mathcal{O})$, a possible world $s \in S$, a learning method L, and a finite data sequence σ, be given. Sequence σ is called a *locking sequence* for s and L if $s \in \bigcap \text{set}(\sigma)$ and for each data sequence τ with $s \in \bigcap \text{set}(\tau)$, $L(\sigma * \tau) = L(\sigma) = \{s\}$.

LEMMA 1. *If learning method L identifies possible world s in the limit, then there exists a locking sequence for s and L.*

PROOF. Assume L identifies s without there being a locking sequence for L and s. We construct in stages a data stream $\vec{O}$ for s on which L does not

converge. Let $x_1, x_2, x_3, \ldots$ enumerate elements of (the countable) $\mathcal{O}$ that are true in in s.

Stage 1. The string (x_1) is not a locking sequence, so for some τ, and sound data sequence for s, $L((x_1) * \tau) \neq L((x_1))$. Take $(x_1) * \tau$ as the initial segment σ_1 of $\vec{O}$.

Stage $n+1$. Assume that the initial segment σ_n of $\vec{O}$ has been constructed in stage n. By assumption, the sequence $\sigma_n * (x_{n+1})$ is not a locking sequence, so there is a sequence τ sound for s such that $L(\sigma_n * (x_{n+1}) * \tau) \neq L(\sigma_n * (x_{n+1}))$. Let $\sigma_{n+1} = \sigma_n * (x_{n+1}) * \tau$.

Because each x_i occurs in $\vec{O}$, $\vec{O}$ is a sound and complete data stream for s. But learner L keeps changing value on $\vec{O}$, L does not converge. ∎

The characterization of identifiability in the limit requires the existence of finite sets that allow drawing a conclusion without the risk of overgeneralization. The characterization theorem is adapted from [2] and [29].

LEMMA 2. *Let* $\mathbb{S} = (S, \mathcal{O})$ *be an epistemic space.* $\mathbb{S}$ *is identifiable in the limit iff there exists a total map* $D : S \to \mathcal{P}(\mathcal{O})$, *given by* $s \mapsto D_s$, *such that* D_s *is a finite tell-tale for* s, *i.e.,*

(1) D_s *is finite,*

(2) $s \in \bigcap D_s$,

(3) *for any* $t \in S$, *if* $t \in \bigcap D_s$ *and* $\mathcal{O}_t \subseteq \mathcal{O}_s$, *then* $t = s$.

PROOF. [$\Rightarrow$] Let $\mathbb{S} = (S, \mathcal{O})$ be an epistemic space. Recall that S and $\mathcal{O}$ are at most countable. Assume that $\mathbb{S}$ is identifiable in the limit by learning method L, i.e., for every world $s \in S$ and every sound and complete positive data stream for s, there exists a finite stage after which L outputs the singleton $\{s\}$ from then on. By Lemma 1, for every $s \in S$ we can take a locking sequence σ_s for L and s. For any $s \in S$ we define $D_s := \text{set}(\sigma_s)$.

(1) D_s is finite because locking sequences are finite.

(2) $s \in \bigcap D_s$, because $s \in \bigcap \text{set}(\sigma_s)$.

(3) Assume that there are $s, t \in S$, such that $s \neq t$ and for all $p \in \mathcal{O}$ such that $t \in p$ it is the case that $s \in p$. Take a sound and complete data stream $\vec{O}$ for t, such that for some $n \in \mathbb{N}$, $\vec{O}[n] = \sigma_s$. Because σ_s is a locking sequence for L on s and $t \in \bigcap \text{set}(\vec{O})$, L converges to s on $\vec{O}$. Therefore, L does not identify t, a world in S. Contradiction.

[$\Leftarrow$] Assume that $\mathbb{S} = (S, \mathcal{O})$ has tell-tales. Let us enumerate $S = \{s_0, s_1, \ldots\}$ and define L in the following way:

$$L(\mathbb{S}, \sigma) = \{s_i\}, \text{ where } i \text{ is least such that } D_{s_i} \subseteq \text{set}(\sigma),$$
$$\text{if such } i \text{ exists; else } L(\mathbb{S}, \sigma) = \emptyset.$$

Assume that $\vec{O}$ is a sound and complete data stream for s, and that i is the least index for s in the enumeration of S. It suffices to show that, for k large enough, $L(\mathbb{S}, \vec{O}[k]) = \{s_i\}$. We can fix n large enough so that $D_{s_i} \subseteq \text{set}(\vec{O}[n])$. Nevertheless, we cannot conclude that $L(\mathbb{S}, \vec{O}[n]) = \{s_i\}$, because there may be (finitely many) other worlds $s_1, \ldots, s_m$, with $m < i$ distinct from s, that satisfy the same condition. Take any such s_j, $j \in \{0, \ldots, m\}$. So $D_{s_j} \subseteq \text{set}(\vec{O}[n]) \subseteq \mathcal{O}_s$, but then, by the properties of tale-tales, there is a $p \in \mathcal{O}$ such that $s \in p$ and $s_j \notin p$. As $\vec{O}$ is sound and complete w.r.t. s, there exists k_j such that $p \in \text{set}(\vec{O}[k_j])$. We now take k to be the maximum of all the k_j and j, and then s satisfies $D_s \subseteq \text{set}(\vec{O}[k]) \subseteq \mathcal{O}_s$, and continues to do so for all $n > k$.

This concludes the proof. ∎

Our aim is to show that belief revision can learn every learnable epistemic space. The next step is to construct a suitable plausibility order from tell-tales, but we introduce one additional condition (see (2) in Definition 23, below).

DEFINITION 23. Let $\mathbb{S} = (S, \mathcal{O})$ be an epistemic space, with an injective map $i : S \to \mathbb{N}$, and let D' be a total map, such that $D' : S \to \mathcal{P}(\mathcal{O})$, given by $s \mapsto D'_s$ having the following properties:

(1) D'_s is a finite tell-tale for s;

(2) if $t \in \bigcap D'_s$, but $\mathcal{O}_s \nsubseteq \mathcal{O}_t$, then $i(s) < i(t)$.

We call D' an *ordering tell-tale map*, and D'_s an *ordering tell-tale set* of s.

LEMMA 3. *Let $\mathbb{S} = (S, \mathcal{O})$ be an epistemic space. If $\mathbb{S}$ is identifiable in the limit, then $\mathbb{S}$ has an ordering tell-tale map.*

PROOF. First assume that $\mathbb{S} = (S, \mathcal{O})$ is an epistemic space that is identifiable in the limit. Let $i : S \to \mathbb{N}$ and $j : \mathcal{O} \to \mathbb{N}$ be injective maps. By Lemma 2, there exists map D that assigns a tell-tale to each $s \in S$. On the basis of D, we construct a new map $D' : S \to \mathcal{P}(\mathcal{O})$. We proceed step by step according to the enumeration of S given by i and the enumeration of $\mathcal{O}$ given by j (when $i(s) = n$ we will simply write s_n and similarly for j and $p \in \mathcal{O}$). The general idea is to add data to D_s until the at most finitely many counterexamples to condition (2) in Definition 23 are eliminated.

(1) For s_1 set $D'_{s_1} := D_{s_1}$.

(2) For s_n proceed in the following way. First, for every $k < n$ define:

$$P^n_k := \begin{cases} \{p_\ell \mid \ell \text{ is smallest s.t. } s_n \in p_\ell \text{ and } s_k \notin p_\ell\} & \text{if } s_k \in D_{s_n}, \\ \emptyset & \text{otherwise.} \end{cases}$$

Finally, set $D'_{s_n} := D_{s_n} \cup (P^n_1 \cup \ldots \cup P^n_{n-1})$.

We now verify that D' satisfies Definition 23.

(1) D'_s is finite, because D_s is finite, $i(s) = n$ for some $n \in \mathbb{N}$, and there are only finitely many P^n_k such that $k < n$, each of them being either a singleton or the empty set.

(2) $s_n \in \bigcap D'_{s_n}$, because $s_n \in \bigcap D_{s_n}$ and $s_n \in \bigcap(P^n_1 \cup \ldots \cup P^n_{n-1})$.

(3) For any $s \in S$, if $D'_s \subseteq \mathcal{O}_t \subseteq \mathcal{O}_s$, then $D_s \subseteq D'_s \subseteq \mathcal{O}_t \subseteq \mathcal{O}_s$, and hence, by the definition the finite tell-tale set $t = s$.

It remains to check condition 2: if $t \in \bigcap D'_s$ and $\mathcal{O}_s \nsubseteq \mathcal{O}_t$ then $i(s) < i(t)$. Towards contradiction, assume that $t \in \bigcap D'_s$, $\mathcal{O}_s \nsubseteq \mathcal{O}_t$, and $i(s) \geq i(t)$. If $i(s) = i(t)$, then $\mathcal{O}_s = \mathcal{O}_t$, contradiction. If $i(s) > i(t)$, then by the construction of D'_s, there is a $p \in D'_s$ such that $s \in p$ and $t \notin p$ (if $D_s \subseteq \mathcal{O}_t$ we added such p in the process of obtaining D'_s; otherwise it had been already there to start with). But then $t \notin \bigcap D'_s$. Contradiction. ∎

The next step is to use the ordering tell-tales to define a preorder on an epistemic space.

DEFINITION 24. For $s, t \in S$, define:

$$s \preceq^1_{D'} t \text{ iff } s \in \bigcap D'_t.$$

We take $\preceq_{D'}$ to be the transitive closure of the relation $\preceq^1_{D'}$.

We want to show that, indeed, the above construction generates an order, i.e., that $\preceq_{D'}$ is reflexive, transitive, and antisymmetric. The latter involves proving that $\preceq_{D'}$ includes no proper cycles (see Figure 5).

DEFINITION 25. A *proper cycle* in $\preceq_{D'}$ is a sequence of distinct worlds $s_1, \ldots, s_n$, with $n \geq 2$, and such that:

(1) for all $i = 1, \ldots, n-1$, $s_{i+1} \in \bigcap D'_{s_i}$,

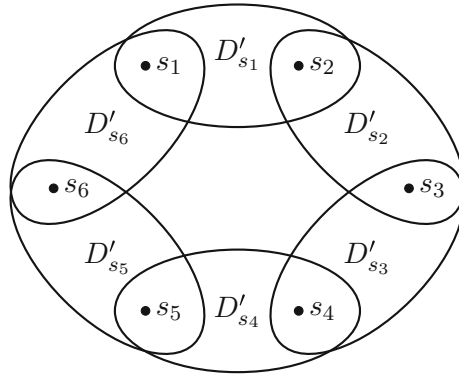(2) $s_1 \in \bigcap D'_{s_n}$.

Figure 5. A visualization of a proper cycle of length 6, here we write D'_{s_i} instead of $\bigcap D'_{s_i}$

LEMMA 4. *For any identifiable epistemic space $\mathbb{S}$ and any ordering tell-tale map D', the relation $\preceq_{D'}$ is an order, i.e., $\preceq_{D'}$ is reflexive, transitive, and antisymmetric.*

PROOF. The fact that $\preceq_{D'}$ is a preorder is trivial: reflexivity follows from the fact that s is always in $\bigcap D'_s$, and transitivity is imposed by construction (by taking the transitive closure).

We need to prove that $\preceq_{D'}$ is antisymmetric. In order to do that we will show (by induction on n) that $\preceq_{D'}$ does not contain proper cycles of any length $n \geq 2$.

(1) For the initial step ($n = 2$): Suppose we have a proper cycle of length 2. As we saw, this means that there exist two states s_1, s_2 such that $s_1 \neq s_2$, $s_2 \in \bigcap D'_{s_1}$, and $s_1 \in \bigcap D'_{s_2}$.

   (a) If $\mathcal{O}_{s_1} \subseteq \mathcal{O}_{s_2}$, then since $s_1 \in \bigcap D'_{s_2}$, so $s_1 = s_2$, (by the fact that D'_{s_2} is a tell-tale for s_2). Similarly, if $\mathcal{O}_{s_2} \subseteq \mathcal{O}_{s_1}$, then since $s_2 \in \bigcap D'_{s_1}$, we again have that $s_2 = s_1$. Contradiction.

   (b) $\mathcal{O}_{s_1} \nsubseteq \mathcal{O}_{s_2}$ and $\mathcal{O}_{s_2} \nsubseteq \mathcal{O}_{s_1}$. From the assumption that $s_2 \in \bigcap D'_{s_1}$, and that $\mathcal{O}_{s_1} \nsubseteq \mathcal{O}_{s_2}$, we can infer (by Condition 2 of Definition 23), that $i(s_1) < i(s_2)$. But, in the same way (from $s_1 \in \bigcap D'_{s_2}$, and $\mathcal{O}_{s_2} \nsubseteq \mathcal{O}_{s_1}$), we can also infer that $i(s_2) < i(s_1)$. Putting these together, we get $i(s_1) < i(s_2) < i(s_1)$. Contradiction.

(2) For the inductive step ($n + 1$): Suppose that there is no proper cycle of length n, and, towards contradiction, that $s_1, s_2, ..., s_{n+1}$ is a proper cycle of length $n + 1$. We consider two cases:

Case 1: There exists k with $1 \leq k \leq n$ such that $\mathcal{O}_{s_k} \subseteq \mathcal{O}_{s_{k+1}}$. Since $s_k \in \cap D'_{s_{k-1}}$, it must be that $s_{k+1} \in \cap D'_{s_{k-1}}$. Therefore, the sequence $s_1, \ldots, s_{k-1}, s_{k+1}, \ldots$ (obtained by deleting s_k from the above proper cycle of length $n+1$) is also a (shorter) proper cycle (of length n). Contradiction.

Case 2: $\mathcal{O}_{s_k} \nsubseteq \mathcal{O}_{s_{k+1}}$, for all $1 \leq k \leq n$. So, for all $1 \leq k \leq n$, $s_{k+1} \in \bigcap D'_{s_k}$. By Condition 2 of Definition 23, it follows that $i(s_k) < i(s_{k+1})$, for all $k = 1, \ldots, n$, and hence $i(s_1) < i(s_{n+1})$. But $s_1 \in \bigcap D'_{s_n}$ and $\mathcal{O}_{s_{n+1}} \nsubseteq \mathcal{O}_{s_1}$ (since otherwise $s_1 \in \cap D'_{s_n}$ and $s_1, \ldots, s_n$ would give a proper cycle of length n), hence $i_{s_1} > i_{s_{n+1}}$. Contradiction. ∎

Let us now show that $\preceq_{D'}$, when used by the conditioning revision method, guarantees convergence to the right belief whenever the underlying epistemic space is identifiable in the limit.

THEOREM 1. *The conditioning belief revision method* (Cond) *is universal.*

PROOF. Obviously, if $\mathbb{S}$ is identifiable in the limit by conditioning, then $\mathbb{S}$ is identifiable in the limit. We therefore focus on the other direction, i.e., we show that if $\mathbb{S}$ is identifiable in the limit by any learning method, then it is identifiable in the limit by conditioning.

By Lemma 3 we know there exists an ordering tell-tale map for $\mathbb{S}$ and by Lemma 4, the corresponding $\preceq_{D'}$ is a (partial) order on S. By the Order-Extension Principle, there exists a total order $\preceq$ on S such that, for all $s, t \in S$, it is the case that $s \preceq_{D'} t$ implies $s \preceq t$.[9]

It remains to show that $\mathbb{S}$ is identifiable in the limit by the learning method generated from the conditioning belief revision method and the prior plausibility assignment $\preceq$. Let $\mathbb{B}_{\mathbb{S}} = (S, \mathcal{O}, \preceq)$ and let us take any $s \in S$ and the corresponding D'_s. Since $s \in \bigcap D'_s$, it follows that for every sound and complete positive data stream $\vec{O}$ for s, there exists $n \in \mathbb{N}$ such that $D'_s \subseteq \text{set}(\vec{O}[n])$. Let $\text{Cond}(\mathbb{B}_{\mathbb{S}}, \vec{O}[n]) = (S', \mathcal{O}, \preceq')$. Our aim is now to demonstrate that $\min_{\preceq'} S' = \{s\}$. By the antisymmetry of the order relation $\preceq$ and hence also of $\preceq'$, the minimal element of S' is unique, so it is sufficient to show that $s \in \min_{\preceq} S'$. For this, let $t \in S'$ be arbitrary. We need to show that $s \preceq t$. Since $t \in S'$, we get that $D'_s \subseteq \text{set}(\vec{O}[n]) \subseteq \mathcal{O}_t$, so, by Definition 23, $s \preceq_{D'} t$, and hence $s \preceq t$. It remains to show that Cond stabilizes on $\{s\}$. Observe that $\vec{O}$ is sound with respect to s, and therefore no further

---

[9]In general, the proof of this principle uses the Axiom of Choice. But here we only need the special case in which the support set S is *countable*, and this special case is provable without the Axiom of Choice.
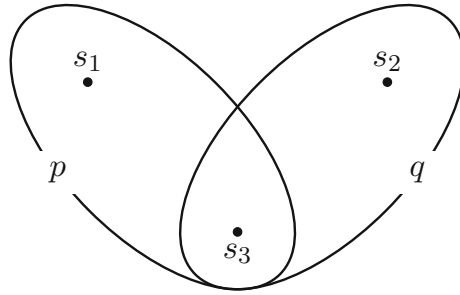
Figure 6. Epistemic space from the proof of Theorem 3. It is impossible
to find a plausibility order that would allow learnability via minimal
belief revision method, `Mini`

information from $\vec{O}$ can eliminate s (because conditioning is conservative),
and hence for any future data the set of minimal elements will remain $\{s\}$.
∎

THEOREM 2. *The lexicographic belief revision method* (`Lex`) *is universal.*

The proof is analogous to the proof of Theorem 1. Within our learning
setting lexicographic revision with true information does exactly what con-
ditioning does. The only difference is that the rest of the doxastic structure
might not stabilize, but only the minimal elements stabilize.

THEOREM 3. *The minimal belief revision method* (`Mini`) *is not universal.*

PROOF. Let us give a counter-example, an epistemic space that is identi-
fiable in the limit, but is not identifiable by the minimal revision method
(see Figure 6). Let $\mathbb{S} = (S, \mathcal{O})$, where $S = \{s_1, s_2, s_3\}$, $\mathcal{O} = \{p, q\}$, and
$p = \{s_1, s_3\}$, $q = \{s_2, s_3\}$. The epistemic space $\mathbb{S}$ is identifiable in the limit
by the conditioning revision method: just assume the ordering $s_1 \prec s_2 \prec s_3$.
However, there is no ordering that allows identification in the limit of S
by the minimal revision method. If s_3 occurs in the ordering before s_1 (or
before s_2), then the minimal revision method fails to identify s_1 (s_2, respec-
tively). If both s_1 and s_2 precede s_3 in the ordering then the minimal revision
method fails to identify s_3 on any data stream consisting of singletons of
propositions from s_3. On all such data streams for s_3 the minimal state will
alternate between s_1 and s_2, or stabilize on one of them. The last case is
that at least one of s_1 and s_2 is equi-plausibile to s_3. In such case s_3 is not
identifiable because for any single proposition from s_3 there is more than
one possible world consistent with it.   ∎
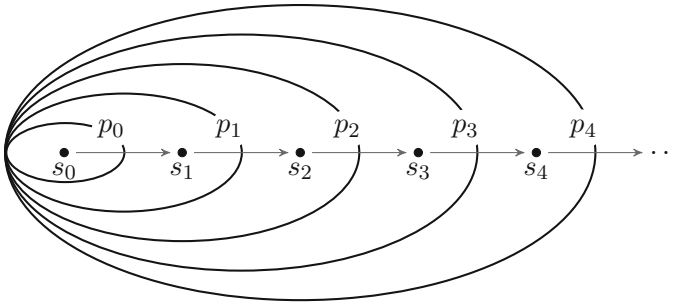
Figure 7. Epistemic space from the proof of Theorem 4. The grey arrows show the non-well-founded plausibility order appropriate for successful, non-standard learning

THEOREM 4. *No conservative belief revision method is standardly universal.*

PROOF. There is an epistemic space $\mathbb{S}$ that is identifiable in the limit by a learning method, but is not standardly identified in the limit by any conservative belief revision method. The following epistemic space constitutes such a counter-example. Let $\mathbb{S} = (S, \mathcal{O})$ such that $S = \{s_n \mid n \in \mathbb{N}\}$, $\mathcal{O} = \{p_i \mid i \in \mathbb{N}\}$, and for any $k \in \mathbb{N}$, $p_k = \{s_i \mid 0 \leq i \leq k\}$, see Figure 7.

$\mathbb{S}$ is identifiable in the limit[10] by the following learning method L:

$$L(\mathbb{S}, \sigma) = \{s_n\} \text{ iff } n \text{ is the largest such that } s_n \in \bigcap \text{set}(\sigma).$$

Let us now assume (towards contradiction) that $\mathbb{S}$ is standardly identifiable in the limit by a conservative belief revision method R, i.e., there exists a well-founded total preorder $\preceq$ on S, such that the learning method $L_R^{\preceq}$ generated from R and $\preceq$ identifies S in the limit and is conservative.

If $\preceq$ is well-founded we can choose some minimal $s_k \in \min_{\preceq} S$ and set $L_R^{\preceq}(\mathbb{S}, \lambda) = \{s_k\}$, where λ is the empty data sequence. Take now some $m > k$, and notice that $\mathcal{O}_{s_m} \subset \mathcal{O}_{s_k}$ (by our construction of $\mathbb{S}$). Let $\vec{O}$ be a sound and complete data stream for s_m. By assumption, $L_R^{\preceq}$ identifies s_m in the limit, hence there must exists some k such that $L_R^{\preceq}(S, \vec{O}[k]) = \{s_m\}$. But since $\mathcal{O}_{s_m} \subseteq \mathcal{O}_{s_k}$, the stream $\vec{O}$ is sound for s_k as well: $\text{set}(\vec{O}[n]) \subseteq \mathcal{O}_{s_k}$, for all $n \in \mathbb{N}$.

We prove by induction that $s_k \in L_R^{\preceq}(\mathbb{S}, \vec{O}[n])$ for all $n \in \mathbb{N}$. Note that this leads to contradiction, namely to $s_k \in L_R^{\preceq}(\mathbb{S}, \vec{O}[n]) = \{s_m\}$, and hence

---

[10]And, as a consequence of our previous results, it is identifiable in the limit by conditioning. Indeed, it is enough take the prior plausibility given by: $s_n \preceq s_m$ iff $n \geq m$. But notice that $\preceq$ is not well-founded, so this is not a standard prior.

to $s_k = s_m$, which contradicts our choice of $m > k$. The proof by induction goes as follows. The base case is already established, since $s_k \in L(\mathbb{S}, \lambda) = L_R^{\preceq}(\mathbb{S}, \vec{O}[0])$. For the inductive case, let us assume that $s_k \in L_R^{\preceq}(\mathbb{S}, \vec{O}[n])$ for some n, then $\text{set}(\vec{O}[n+1]) \subseteq \mathcal{O}_{s_k}$ and L is conservative to conclude that $s_k \in L_R^{\preceq}(\mathbb{S}, \vec{O}[n] * \vec{O}_{n+1}) = L_R^{\preceq}(\mathbb{S}, \vec{O}[n+1])$.   ■

The above result concerns the type of preorders that facilitate identifiability in the limit. We show that for universality results our non-standard setting (involving non-well-founded plausibility orders) is essential: assuming that AGM-like belief revision must be conservative, no such method is universal with respect to well-founded plausibility spaces.

COROLLARY 1. *No AGM-like belief revision method is standardly universal.*

5.   Learning from Positive and Negative Data

One may wonder what would happen if the revision process was governed not only by arbitrary sets of observables, but by observables which are closed under certain logical operations. One simple adjustment of the set $\mathcal{O}$ is to assume its closure on negation.[11]

First let us extend our framework to account for situations in which both positive and negative data can be observed.

DEFINITION 26. An epistemic space $\mathbb{S} = (S, \mathcal{O})$ is *negation-closed* if the set $\mathcal{O}$ of all data is negation-closed, i.e., if for every $p \in \mathcal{O}$ there exists some $\overline{p} \in \mathcal{O}$ such that $p = S \setminus \overline{p}$ (i.e., for every $s \in S$, $s \in \overline{p}$ iff $s \notin p$).

PROPOSITION 10. *Conditioning and lexicographic revision generate standardly universal learning methods on the class of negation-closed epistemic spaces.*

PROOF. We prove that every negation-closed epistemic space $\mathbb{S}$ (with $\mathcal{O}$ and S countable) is identifiable in the limit by conditioning and by lexicographic revision.

Let us assume that $\mathbb{S}$ is countable and negation closed. In fact, any ω-type order $\preceq$ on S gives a suitable (well-founded) prior plausibility assignment. Let us take an $s \in S$. Since $\preceq$ is ω-type it is well-founded, so there are only finitely many worlds that are more plausible than s. For each such world $t \prec s$ we collect one $O^t \in \mathcal{O}$ such that $s \in O^t$ but $t \notin O^t$ or *vice versa*. Such

---

[11]In a follow up work we will consider closure on finite intersections, which will allow us to view learnability in epistemic spaces as topological properties [see 4,6].
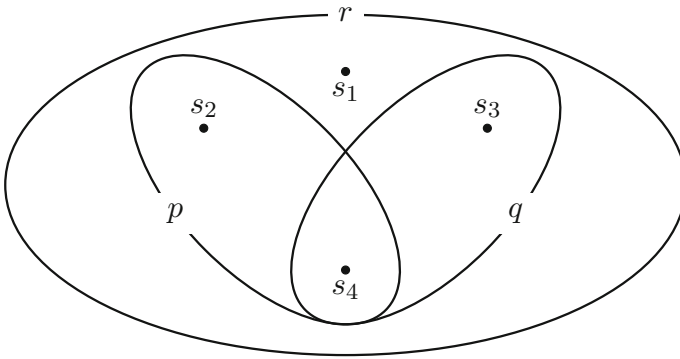
Figure 8. Epistemic space from the proof of Proposition 11

an O^t must exist, since $s \neq t$ implies that either there exists some $O \in \mathcal{O}$ such that $s \in O$ and $t \notin O$ or there exists some $O' \in \mathcal{O}$ such that $t \in O'$ and $s \notin O'$. In the first case, we put $O^t = O$, while in the second case we put $O^t = \overline{O}$.) Then the data set $\{O^t \mid t \prec s\}$ is finite. For every data stream $\vec{O}$ that is sound and complete with respect to s, there must exist a stage $n \in \mathbb{N}$ by which all data in $\{O^t \mid t \prec s\}$ have been observed. After this stage, all worlds that are more plausible than s will have been deleted (in the case of conditioning) or will have become less plausible than s (in the case of lexicographic revision), so from then on the (only) most plausible state is s. Hence conditioning and lexicographic revision identify any world $s \in S$ in the limit. ∎

PROPOSITION 11. *Minimal revision is not universal on negation-closed spaces.*

PROOF. We show a negation-closed epistemic space that is identifiable in the limit, but is not identifiable in the limit by the minimal revision method. We take $\mathbb{S} = (S, \mathcal{O})$, where: $S = \{s_1, s_2, s_3, s_4\}$ and $\mathcal{O} = \{p, q, r\}$, with $p = \{s_2, s_4\}$, $q = \{s_3, s_4\}$, and $r = \{s_1, s_2, s_3, s_4\}$, see Figure 8.

The epistemic space $\mathbb{S}$ is identifiable in the limit by `Cond`, just assume the plausibility order $s_1 \prec s_2 \prec s_3 \prec s_4$. However, there is no plausibility order that allows identification in the limit of $\mathbb{S}$ by the minimal revision method. Whichever ordering is assumed, the least plausible element will not be identifiable. It is so because each piece of data consistent with s is also consistent with one of the $\prec$-smaller sets. ∎

6.   Erroneous Information

With the introduction of negative information, we can now allow for occasional observational *errors*, and for their corrections. To consider erroneous data we now give up the soundness of data streams, i.e., we allow that the learner can observe data that may be false in the real world. In order to still give the agent a chance to learn the real world, we need to impose some limitation on errors. We do this by requiring the data streams to be 'fair'.[12]

DEFINITION 27. Let $\mathbb{S} = (S, \mathcal{O})$ be a negation-closed epistemic space. A stream $\vec{O}$ of data from $\mathcal{O}$ is *fair* with respect to the world s if $\vec{O}$ *contains only finitely many errors and every such error is eventually corrected* in $\vec{O}$, in other words:

(1) $\vec{O}$ is complete with respect to s,

(2) there is $n \in \mathbb{N}$ such that for all $k \geq n$, $s \in O_k$, and

(3) for every $i \in \mathbb{N}$ such that $s \notin O_i$, for some $k > i$, $O_k = \overline{O_i}$.

Unsurprisingly, conditioning (which assumes absolute veracity of the new observations) is no longer a good strategy. If erroneous observations are possible, then eliminating worlds that do not satisfy these observations is risky and irrational.

PROPOSITION 12. *Conditioning and minimal revision are not universal for fair streams.*

PROOF. Conditioning does not tolerate errors at all. On any $\vec{O}_i$ such that $s \notin \vec{O}_i$ conditioning will remove s and there is no way to revive it. Minimal revision, as it has been shown, is not universal on negation-closed epistemic spaces even with respect to sound and complete data streams, which are a special case of fair streams.   ∎

We will demonstrate that lexicographic revision deals with errors in a skillful manner. Before we get to that we introduce and discuss the notion of *propositional upgrade* [which is a special case of generalized upgrade, see 12]. Such an upgrade is a transformation of a plausibility space that can be given by any finite sequence of mutually disjoint propositional sentences $x_1, \ldots, x_n$. The corresponding propositional upgrade $(x_1, \ldots, x_n)$ acts on a plausibility space $\mathbb{B}_\mathbb{S} = (S, \mathcal{O}, \preceq)$ by changing the preorder $\preceq$ as follows: all

---

[12]Notions defined in Section 4 (identifiability in the limit, universality, etc.) are similar for fair data streams.

worlds that satisfy x_1 become less plausible than all worlds satisfying x_2, all worlds satisfying x_2 become less plausible than all x_3 worlds, etc., up to the worlds which satisfy x_n. Moreover, for any k such that $1 \leq k \leq n$, among the worlds satisfying x_k the old order $\preceq$ is kept the same. In particular, lexicographic revision is a special case of such propositional upgrade, $(\neg p, p)$.

LEMMA 5. *The class of propositional upgrades is closed under sequential composition.*

PROOF. We need to show that the sequential composition of any two propositional upgrades gives a propositional upgrade. Let us take $X :=$ $(x_1, \ldots, x_n)$ and $Y := (y_1, \ldots, y_m)$. The sequential composition $X * Y$ is equivalent to the following propositional upgrade:

$$(x_1 \wedge y_1, \ldots, x_n \wedge y_1, x_1 \wedge y_2, \ldots, x_n \wedge y_2, \ldots, \ldots, x_1 \wedge y_m, \ldots, x_n \wedge y_m).$$

To show this, take an arbitrary plausibility space $\mathbb{B}_\mathbb{S} = (S, \mathcal{O}, \preceq)$ and upgrade on X and Y successively. First, we upgrade on X, to obtain the new preorder $\preceq^X$, in which all worlds satisfying x_1 are less plausible than all x_2-worlds, etc., and within each such partition the old order $\preceq$ is kept the same. Now, to this new plausibility space we apply the second upgrade, Y, obtaining the new preorder $\preceq^{XY}$, in which all y_1-worlds are less plausible than all y_2-worlds, etc. However, since the upgrade Y has been applied to the preorder $\preceq^X$ we also know that the new preorder $\preceq^{XY}$ has the following property: for each j, such that $1 \leq j \leq m$, within the partition given by y_j, all x_1-worlds are less plausible than all x_2-worlds, etc. At the same time, in each j and k such that $1 \leq j \leq m$ and $1 \leq k \leq n$ the preorder $\preceq$ is maintained in the partition $(y_j \wedge x_k)$. Thus, $\preceq^{XY}$ has the following structure:

$$(x_1 \cap y_1) \succeq^{XY} \ldots \succeq^{XY} (x_n \cap y_1) \succeq^{XY}$$
$$(x_1 \cap y_2) \succeq^{XY} \ldots \succeq^{XY} (x_n \cap y_2) \succeq^{XY} \ldots \succeq^{XY} \ldots \succeq^{XY} (x_n \cap y_m),$$

Moreover, within each such partition, the old preorder $\preceq$ is kept the same.

The final observation is that the above setting can be obtained directly by the propositional upgrade of the following form:

$$(x_1 \wedge y_1, \ldots, x_n \wedge y_1, x_1 \wedge y_2, \ldots, x_n \wedge y_2, \ldots, x_1 \wedge y_m, \ldots, x_n \wedge y_m).$$

∎

Now we are ready to show that lexicographic revision is well-behaved on fair streams.

PROPOSITION 13. *Lexicographic revision generates a standardly universal belief revision-based learning method for fair streams on the class of negation-closed epistemic spaces.*

PROOF. First, recall that lexicographic revision, Lex, is standardly universal for sound and complete streams on negation-closed spaces. It is left to show that Lex retains its power on fair streams. It is sufficient to show that lexicographic revision is 'error-correcting', i.e., that the effect of revising with the stream $(p * \sigma * \overline{p})$ is exactly the same as with the stream $(\sigma * \overline{p})$, where σ is any sequence of observables. The proof uses the properties of sequential composition for propositional upgrade.

Let us assume that length$(\sigma) = n$. In terms of generalized upgrade we need to demonstrate that the sequential composition

$$(\neg p, p)(\neg \sigma_1, \sigma_1) \ldots (\neg \sigma_n, \sigma_n)(p, \neg p)$$

is equivalent to

$$(\neg \sigma_1, \sigma_1) \ldots (\neg \sigma_n, \sigma_n)(p, \neg p).$$

From Lemma 5 we know that propositional upgrade is closed under sequential composition. Hence, in the equivalence to be shown, we can replace the composition $(\neg \sigma_1, \sigma_1) \ldots (\neg \sigma_n, \sigma_n)$ by only one generalized upgrade, which we denote by $(x_1, \ldots, x_m)$. Now, we claim that: $(\neg p, p)(x_1, \ldots, x_m)(p, \neg p)$ is equivalent to: $(x_1, \ldots, x_m)(p, \neg p)$.

By the proof of Lemma 5, the composition $(x_1, \ldots, x_n)(p, \neg p)$ has the following form:

$$(x_1 \wedge p, \ldots, x_n \wedge p, x_1 \wedge \neg p, \ldots, x_n \wedge \neg p).$$

Accordingly, the other upgrade, $(\neg p, p)(x_1, \ldots, x_n)(p, \neg p)$, has the following form:

$$(\neg p \wedge x_1 \wedge p, p \wedge x_1 \wedge p, \ldots, \neg p \wedge x_n \wedge p, p \wedge x_n \wedge p, \neg p \wedge x_1 \wedge \neg p, p \wedge x_1$$
$$\wedge \neg p, \ldots, \neg p \wedge x_n \wedge \neg p, p \wedge x_n \wedge \neg p).$$

Let us observe that some of the terms in the above upgrade are inconsistent. We can eliminate them since they correspond to empty subsets of the plausibility space. We obtain:

$$(x_1 \wedge p, \ldots, x_n \wedge p, x_1 \wedge \neg p, \ldots, x_n \wedge \neg p).$$

The observation that the two propositional upgrades turn out to be the same concludes the proof.   ∎

7. Conclusions and Perspectives

We have considered iterated belief revision policies of conditioning, lexicographic, and minimal belief revision. We have identified certain features of

those methods relevant in the context of iterated revision: data-retention and conservativity turn out to be especially important. We defined learning methods based on those revision policies and have shown how the aforementioned properties influence the learning process. Throughout the paper we have been mainly interested in convergence to the actual world on the basis of infinite data streams. In the setting of positive, sound and complete data streams we have exhibited that conditioning and lexicographic revision generate universal learning methods. Minimal revision fails to be universal, and the crucial property that makes it weaker is its strong conservativity. Moreover, we have shown that the full power of learning cannot be achieved when the underlying prior plausibility assignment is assumed to be well-founded. In the case of positive and negative information, both conditioning and lexicographic revision are universal. Minimal revision again is not. Finally, in the setting of fair streams (containing a finite number of errors that all get corrected later in the stream) lexicographic revision again turns out to be universal. Both conditioning and minimal revision lack the 'error-correcting' property.

Future and on-going work consists of multi-level investigation of the relationship between formal learning theory, belief revision theory, and DEL. There surely are many links still to be found. What seems to be especially interesting is the multi-agent extension of our results. The interactive aspect would probably be appreciated in formal learning theory, where the single-agent perspective dominates. Another way to extend the framework is to allow revision with more complex formulae. This would link to the AGM approach, and to the philosophical investigation into the process of scientific inquiry, where possible realities have a more 'theoretical' character.

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

[1] ALCHOURRÓN, C. E., P. GÄRDENFORS, and D. MAKINSON, On the logic of theory change: Partial meet contraction and revision functions, *The Journal of Symbolic Logic* 50(2): 510–530, 1985.

[2] ANGLUIN, D., Inductive inference of formal languages from positive data, *Information and Control* 45(2): 117–135, 1980.

[3] AUCHER, G., A combined system for update logic and belief revision, in M. W. Barley, and N. Kasabov, (eds.), *Intelligent Agents and Multi-Agent Systems*, vol. 3371 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2005, pp. 1–17.

[4] BALTAG, A., N. GIERASIMCZUK, A. ÖZGÜN, A. L. VARGAS SANDOVAL, and S. SMETS, A Dynamic Logic for Learning Theory, in A. Madeira, and M. Benevides (eds.), *Dynamic Logic. New Trends and Applications: Proceedings of First International Workshop, DALI 2017, Brasilia, Brazil, September 23–24, 2017*, vol. 10669 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2018, pp. 35–54.

[5] BALTAG, A., N. GIERASIMCZUK, and S. SMETS, Belief revision as a truth-tracking process, in K. Apt, (ed.), *TARK'11: Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge, Groningen, The Netherlands, July 12–14, 2011*, ACM, New York, NY, USA, 2011, pp. 187–190.

[6] BALTAG, A., N. GIERASIMCZUK, and S. SMETS, On the Solvability of Inductive Problems: A Study in Epistemic Topology, in R. Ramanujam, (ed.), *TARK'15: Proceedings of the 15th Conference on Theoretical Aspects of Rationality and Knowledge, Carnegie Mellon University, Pittsburgh, PA, USA, June 4–6, 2015*, Electronic Proceedings in Theoretical Computer Science 215, Open Publishing Association, 2016, pp. 81–98.

[7] BALTAG, A., and L. MOSS, Logics for epistemic programs, *Synthese* 139(2): 165–224, 2004.

[8] BALTAG, A., L. S. MOSS, and S. SOLECKI, The logic of public announcements, common knowledge, and private suspicions, in I. Gilboa, (ed.), *TARK'98: Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998, pp. 43–56.

[9] BALTAG, A., and S. SMETS, Dynamic belief revision over multi-agent plausibility models, in G. Bonanno, W. van der Hoek, and M. Wooldridge, (eds.), *LOFT'06: Proceedings of 7th Conference on Logic and the Foundations of Game and Decision Theory*, University of Liverpool, 2006, pp. 11–24.

[10] BALTAG, A., and S. SMETS, The logic of conditional doxastic actions, in R. van Rooij, and K. Apt, (eds.), *New Perspectives on Games and Interaction. Texts in Logic and Games*, Amsterdam University Press, 2008, pp. 9–31.

[11] BALTAG, A., and S. SMETS, A qualitative theory of dynamic interactive belief revision, in G. Bonanno, W. van der Hoek, and M. Wooldridge, (eds.), *LOFT'08: Proceedings of 8th Conference on Logic and the Foundations of Game and Decision Theory*, no. 3 in Texts in Logic and Games, Amsterdam University Press, 2008, pp. 9–58.

[12] Baltag, A., and S. Smets, Learning by questions and answers: From belief revision cycles to doxastic fixed points, in H. Ono, M. Kanazawa, and R. de Queiroz, (eds.), *Logic, Language, Information and Computation*, vol. 5514 of *Lecture Notes in Computer Science*, Springer, 2009, pp. 124–139.

[13] van Benthem, J., Dynamic logic for belief revision, *Journal of Applied Non-Classical Logics* 2: 129–155, 2007.

[14] van Benthem, J., *Logical Dynamics of Information and Interaction*, Cambridge University Press, Cambridge, MA, USA, 2011.

[15] Blum, L., and M. Blum, Toward a mathematical theory of inductive inference, *Information and Control* 28: 125–155, 1975.

[16] Bolander T., and N. Gierasimczuk, Learning Actions Models: Qualitative Approach, in W. van der Hoek, and W. H. Holliday, (eds.), *LORI'15: Proceedings of 5th International Workshop on Logic, Rationality, and Interaction, Taipei, Taiwan, October 28–31, 2015*, vol. 9394 of *Lecture Notes in Computer Science*, Springer, 2015, pp. 40–52.

[17] Bolander T., and N. Gierasimczuk, Learning to Act: Qualitative Learning of Deterministic Action Models, *Journal of Logic and Computation* 28(2): 337–365, 2017.

[18] Boutilier, C., Iterated revision and minimal change of conditional beliefs, *Journal of Philosophical Logic* 25: 263–305, 1996.

[19] Dégremont, C., and N. Gierasimczuk, Can doxastic agents learn? On the temporal structure of learning, in X. He, J. F. Horty, and E. Pacuit, (eds.), *LORI'09: Proceedings of 2nd International Workshop on Logic, Rationality, and Interaction, Chongqing, China, October 8–11, 2009*, vol. 5834 of *Lecture Notes in Computer Science*, Springer, 2009, pp. 90–104.

[20] Dégremont, C., and N. Gierasimczuk, Finite identification from the viewpoint of epistemic update, *Information and Computation* 209(3): 383–396, 2011.

[21] van Ditmarsch, H., Prolegomena to dynamic logic for belief revision, *Knowledge, Rationality & Action (Synthese)* 147: 229–275 (41–87), 2005.

[22] van Ditmarsch, H., W. Van der Hoek, and B. Kooi, *Dynamic Epistemic Logic*, Springer, The Netherlands, 2007.

[23] Gierasimczuk, N., Bridging learning theory and dynamic epistemic logic, *Synthese* 169(2): 371–384, 2009.

[24] Gierasimczuk, N., Learning by erasing in dynamic epistemic logic, in A. H. Dediu, A. M. Ionescu, and C. Martin-Vide, (eds.), *LATA'09: Proceedings of 3rd International Conference on Language and Automata Theory and Applications, Tarragona, Spain, April 2–8, 2009*, vol. 5457 of *Lecture Notes in Computer Science*, Springer, The Netherlands, 2009, pp. 362–373.

[25] Gierasimczuk, N., *Knowing One's Limits. Logical Analysis of Inductive Inference*, Ph.D. thesis, Universiteit van Amsterdam, The Netherlands, 2010.

[26] Gierasimczuk, N., D. de Jongh, and V. F. Hendricks, Logic and learning, in A. Baltag, and S. Smets, (eds.), *Johan van Benthem on Logical and Informational Dynamics*, Springer, 2014.

[27] Gierasimczuk, N., and D. de Jongh, On the complexity of conclusive update, *The Computer Journal* 56(3): 365–377, 2013.

[28] Gold, E. M., Language identification in the limit, *Information and Control* 10: 447–474, 1967.

[29] Jain, S., D. Osherson, J. S. Royer, and A. Sharma, *Systems that Learn*, MIT Press, Chicago, 1999.

[30] Kelly, K. T., Iterated belief revision, reliability, and inductive amnesia, *Erkenntnis* 50: 11–58, 1998.

[31] Kelly, K. T., The learning power of belief revision, in *TARK'98: Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998, pp. 111–124.

[32] Kelly, K. T., Learning theory and epistemology, in I. Niiniluoto, M. Sintonen, and J. Smolenski, (eds.), *Handbook of Epistemology*, Dordrecht: Kluwer, 2004, pp. 183–203.

[33] Kelly, K. T., Ockham's razor, truth, and information, in P. Adriaans, and J. van Benthem, (eds.), *Handbook of the Philosophy of Information*, Elsevier, 2008, pp. 321–359.

[34] Kelly, K. T., O. Schulte, and V. Hendricks, Reliable belief revision, in *Proceedings of the 10th International Congress of Logic, Methodology, and Philosophy of Science*, Kluwer Academic Publishers, 1995, pp. 383–398.

[35] Martin, E., and D. Osherson, Scientific discovery based on belief revision, *The Journal of Symbolic Logic* 62(4): 1352–1370, 1997.

[36] Martin, E., and D. Osherson, *Elements of Scientific Inquiry*, MIT Press, Cambridge, MA, USA, 1998.

[37] Osherson, D., M. Stob, and S. Weinstein, *Systems that Learn*, MIT Press, Cambridge, MA, USA, 1986.

[38] Rott, H., Conditionals and theory change: Revisions, expansions and additions, *Synthese* 81: 91–113, 1989.

[39] Spohn, W., Ordinal conditional functions: A dynamic theory of epistemic states, in B. Skyrms, and W. L. Harper, (eds.), *Causation in Decision, Belief Change, and Statistics*, vol. II, Dordrecht: Kluwer, 1988, pp. 105–134.

A. Baltag, S. Smets
Institute for Logic, Language, and Computation
University of Amsterdam
P.O. Box 94242 1090 GE Amsterdam
The Netherlands
TheAlexandruBaltag@gmail.com

Sonja Smets
s.j.l.smets@uva.nl

N. GIERASIMCZUK
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Richard Petersens Plads, Building 322, Room 024
2800 Kgs. Lyngby
Denmark
`nigi@dtu.dk`